Centro Euro-Mediterraneo sui Cambiamenti Climatici

www.cmcc.it

AI and climate predictions (with a focus on extreme events)

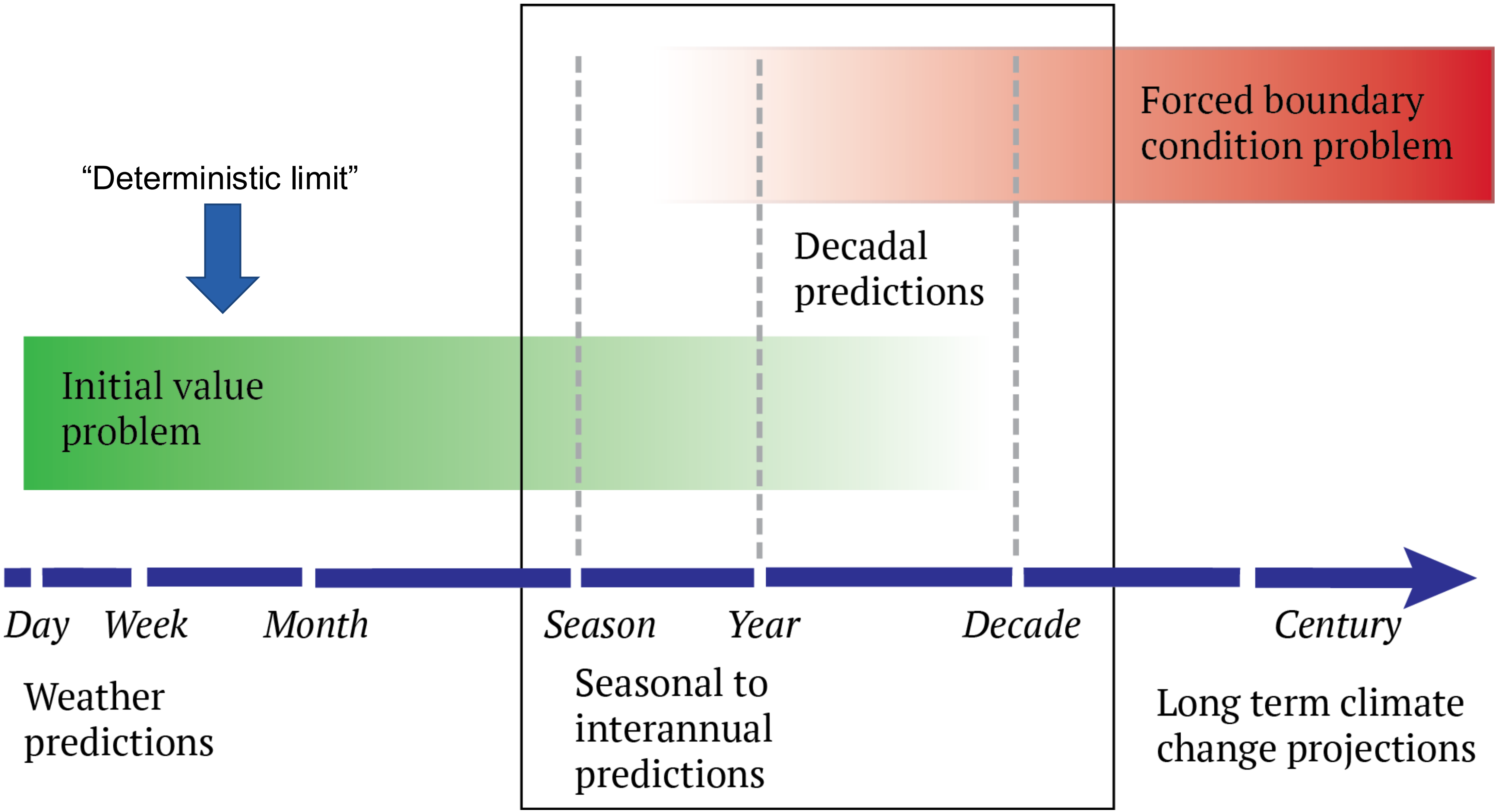
Leone Cavicchia – Fondazione CMCC (Euro Mediterranean Center on Climate Change)

10-9-2025

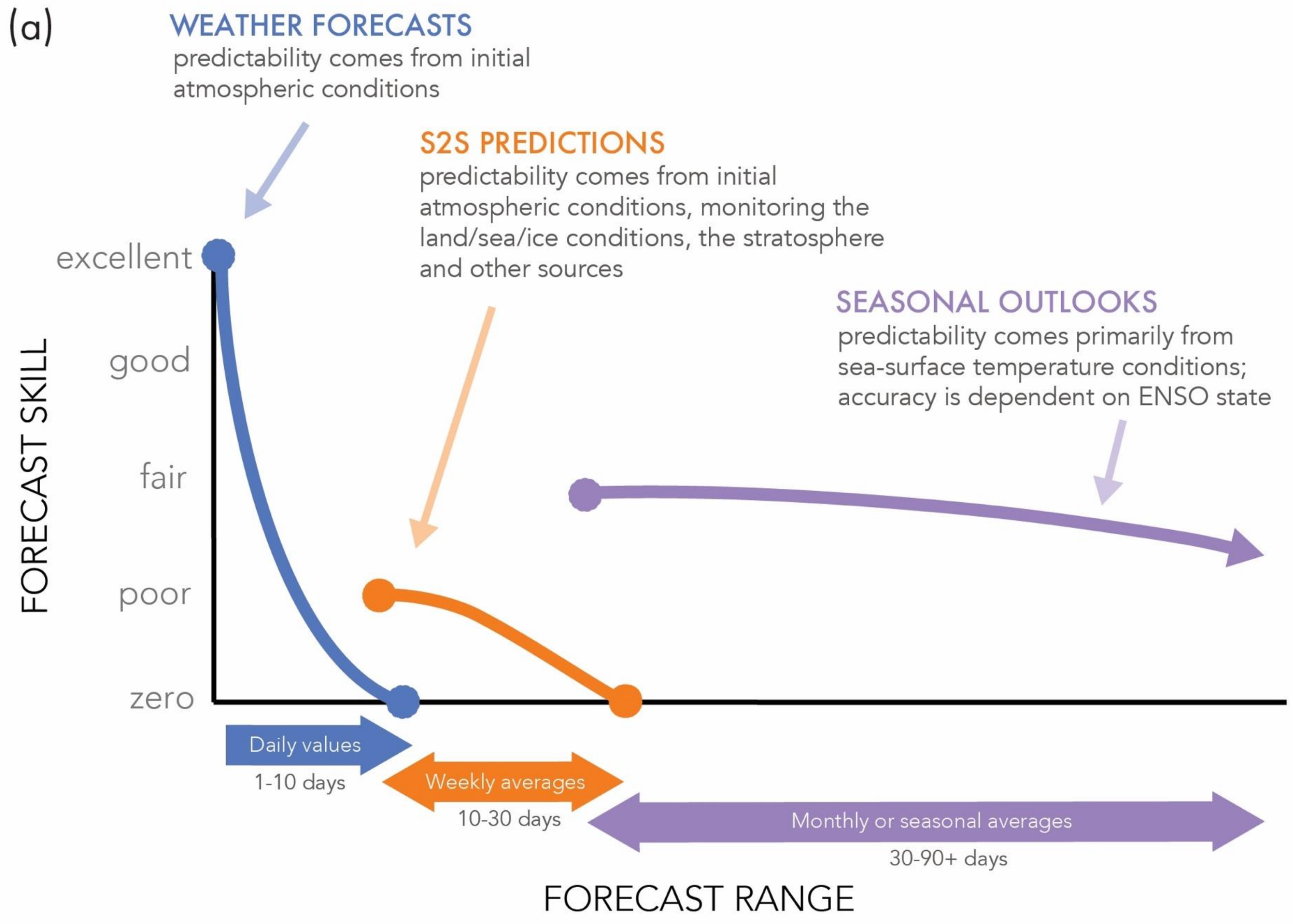
Outline

- *Climate predictions: achievements and challenges*
- *Artificial intelligence and climate predictions*
- *An example: climate prediction of Mediterranean cyclones*

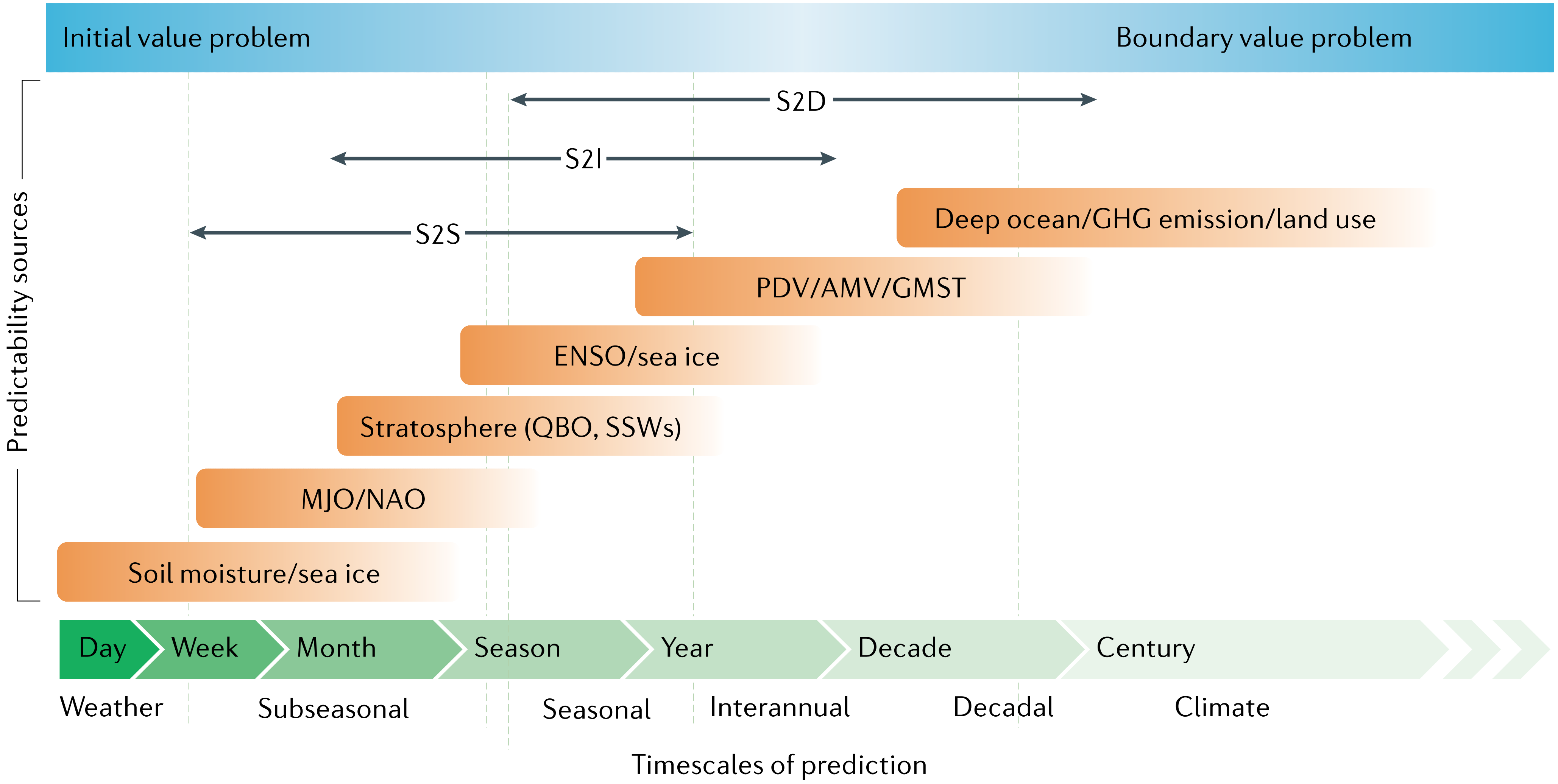
Climate predictions



"Deterministic limit"
Forced boundary condition problem
Decadal predictions
Initial value problem
Day
Week
Month
Season
Year
Decade
Century
Weather predictions
Seasonal to interannual predictions
Long term climate change projections


(a)
WEATHER FORECASTS
predictability comes from initial atmospheric conditions
S2S PREDICTIONS
predictability comes from initial atmospheric conditions, monitoring the land/sea/ice conditions, the stratosphere and other sources
SEASONAL OUTLOOKS
predictability comes primarily from sea-surface temperature conditions; accuracy is dependent on ENSO state
FORECAST SKILL
excellent
good
fair
poor
zero
Daily values
1-10 days
Weekly averages
10-30 days
Monthly or seasonal averages
30-90+ days
FORECAST RANGE

a Predictability sources and timescales

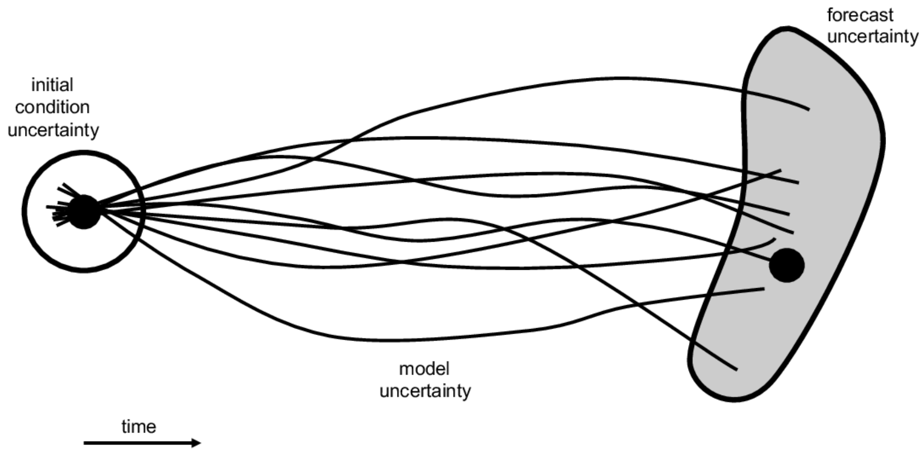


Source: WCRP

Sources of forecast uncertainty

The two main sources of uncertainty in dynamical climate prediction are:

- the lack of perfect knowledge of the initial conditions of the climate system



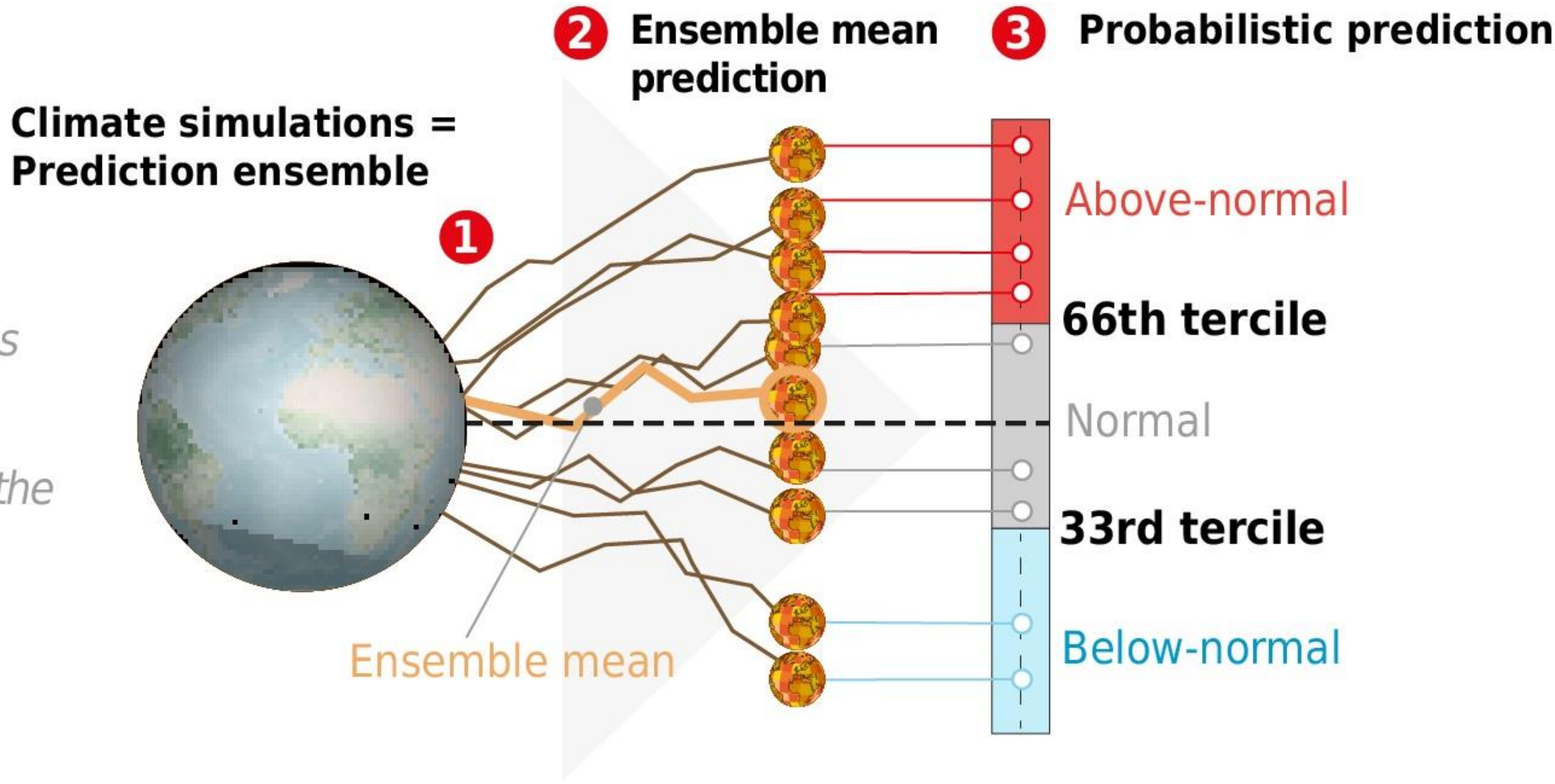
- the inability to perfectly model this system

⇨ parameter pertubation, multimodel approach

Climate predictions are ensemble predictions

▶ *Illustration of ensemble mean prediction and probabilistic prediction:*

1. *Multiple climate simulations (left side) form a prediction ensemble.*
2. *The individual climate predictions are expressed as a deviation (anomaly) from a reference period in the past. The mean of all simulations of the prediction ensemble forms the ensemble mean prediction.*
3. *Dividing the individual climate predictions into the categories “above-normal", “normal", and “below-normal" (separated by the 33rd and 66th terciles of the reference period) leads to the probabilistic prediction (right side).*

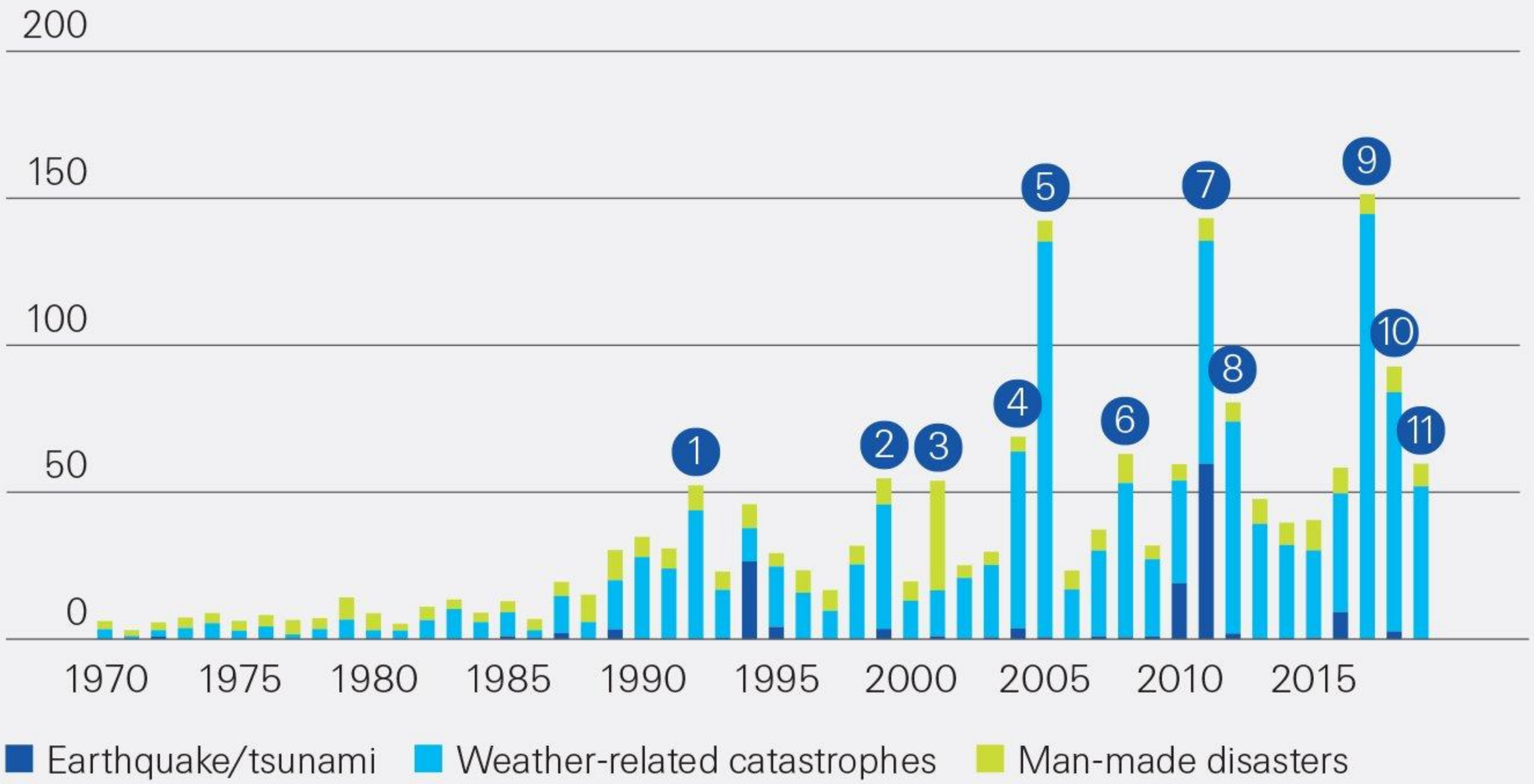


https://www.dwd.de/EN/ourservices/kvhs_en/help/1_bkgrd_info/04_predictions/pics/ensemble.jpg

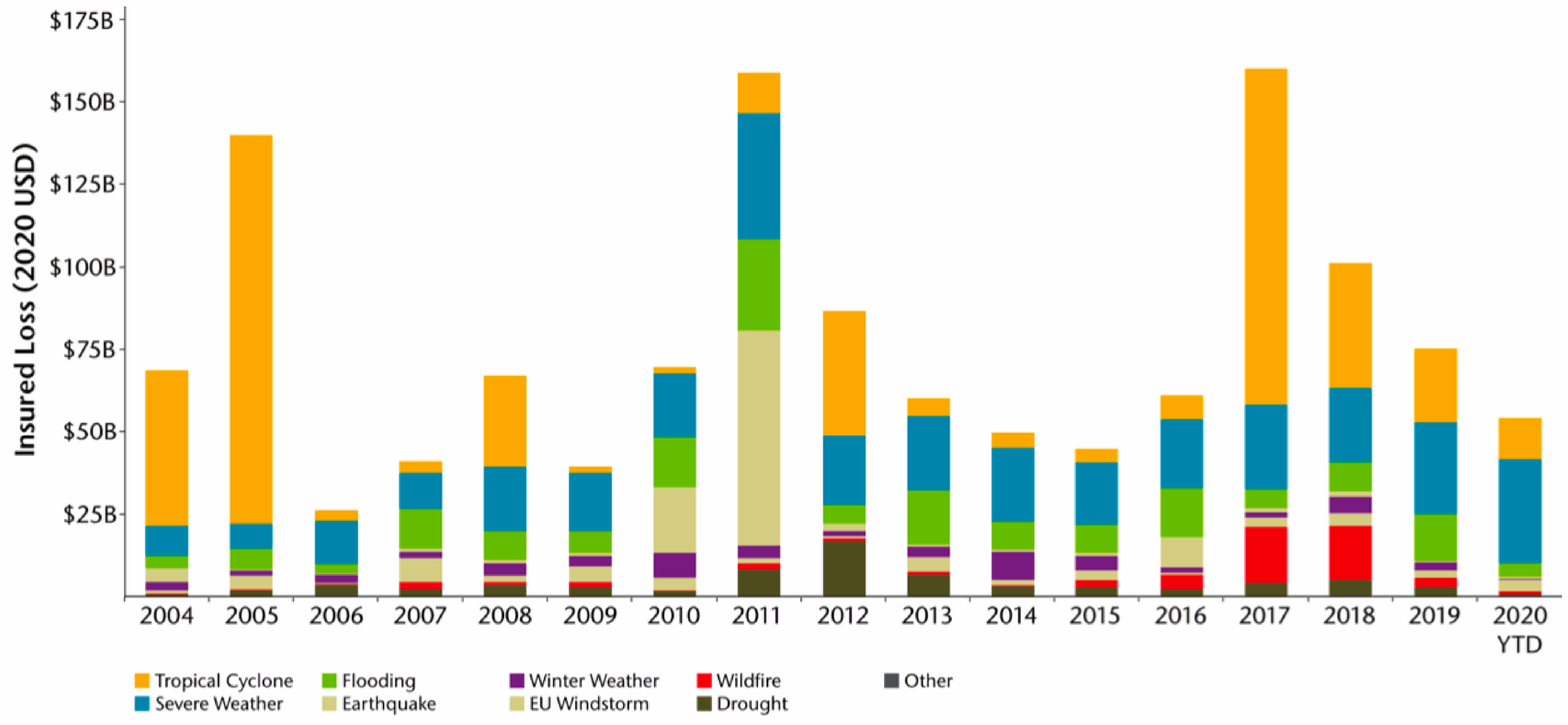
The importance of TC seasonal forecast

Insured catastrophe losses, 1970–2019, in USD billion at 2019 prices

1. Hurricane Andrew
2. Winter Storm Lothar
3. WTC
4. Hurricanes Ivan, Charley, Frances
5. Hurricanes Katrina, Rita, Wilma
6. Hurricanes Ike, Gustav
7. Japan, NZ earthquakes, Thailand flood
8. Hurricane Sandy
9. Hurricanes Harvey, Irma, Maria
10. Camp Fire, Typhoon Jebi
11. Typhoons Hagibis, Faxai



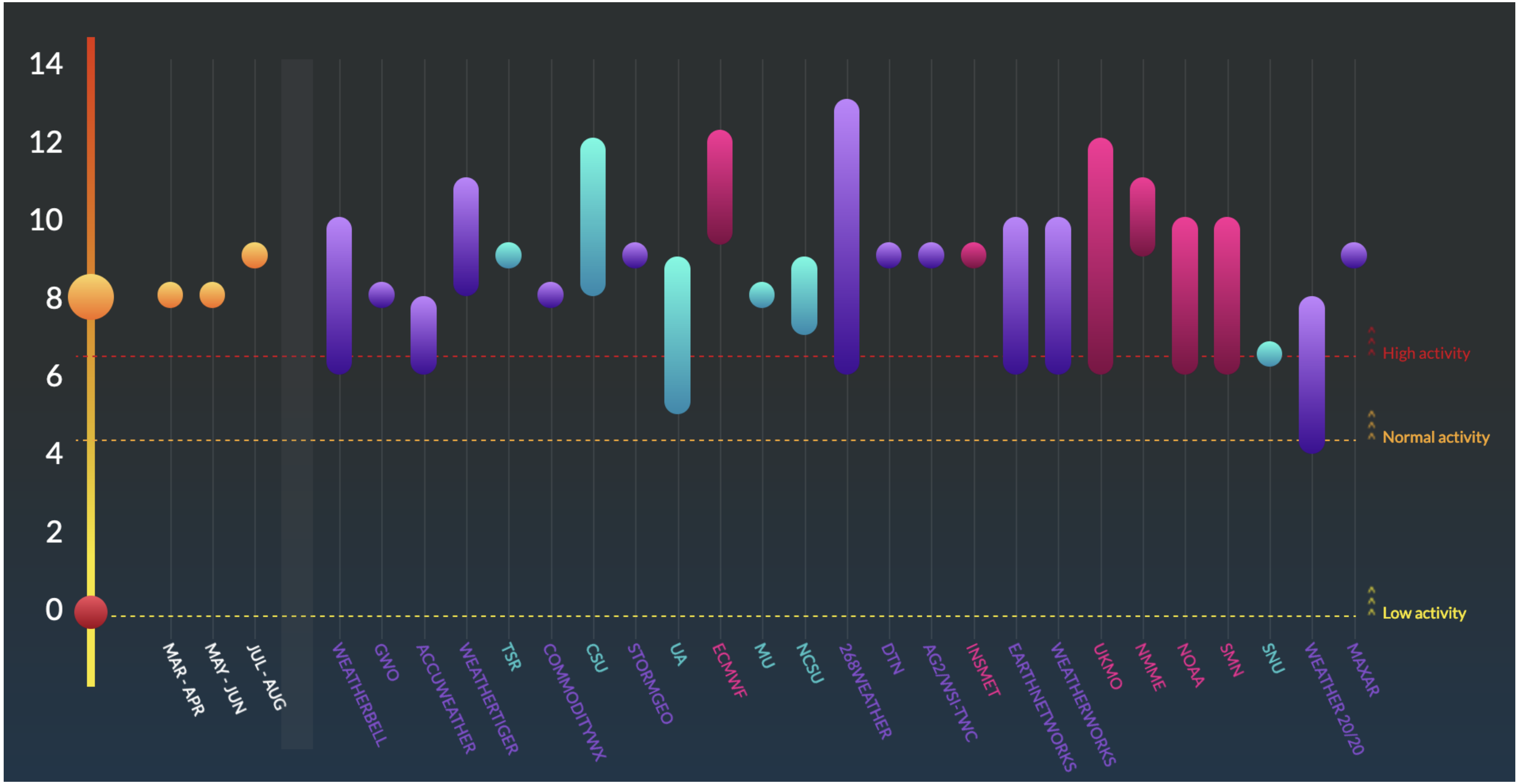
Source: Swiss Re Institute



Source: Aon Catastrophe Insight

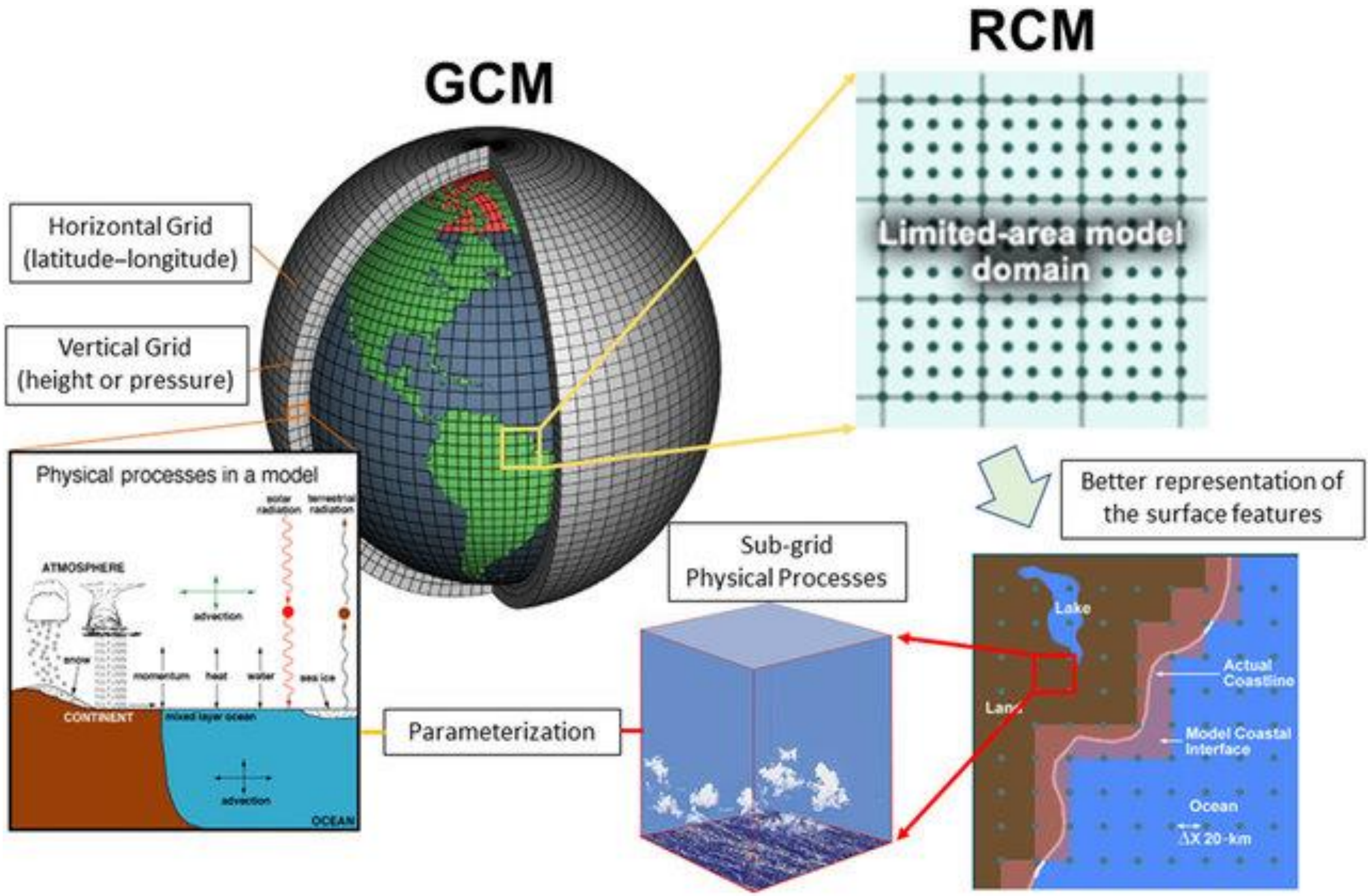
TC Seasonal Predictions

https://seasonalhurricanepredictions.bsc.es/predictions



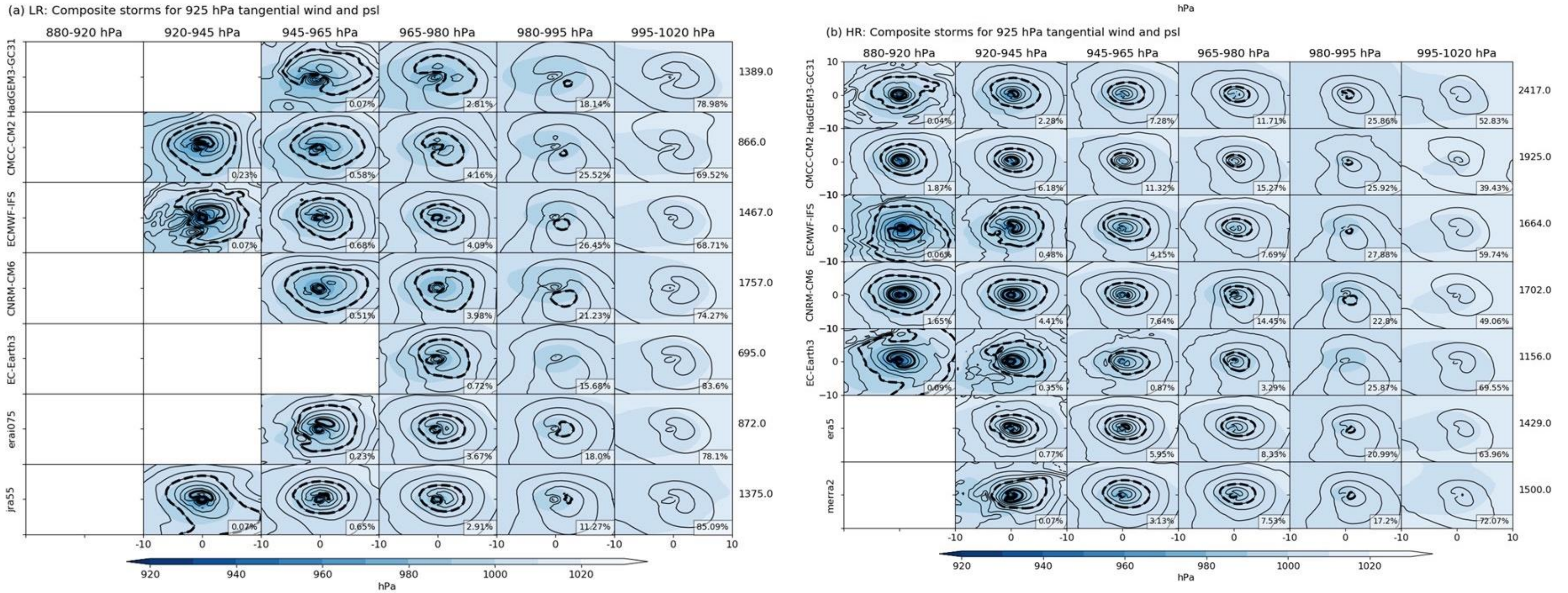
Climate models resolution

- *Climate models numerically approximate fluid dynamics equations on a discrete lat-lon grid.*
- *The resolution of the horizontal grid is determined by the available computer power.*
- *Sub-grid processes are represented by* ***physical parameterizations.***
- *Parameterizations are computationally expensive and depend on a large number of arbitrary parameters.*



➢ *The large number of ensemble members needer to produce skilful climate predictions limits the horizontal resolution to about 50 kms on today's supercomputers.*

Climate model resolution and extreme events



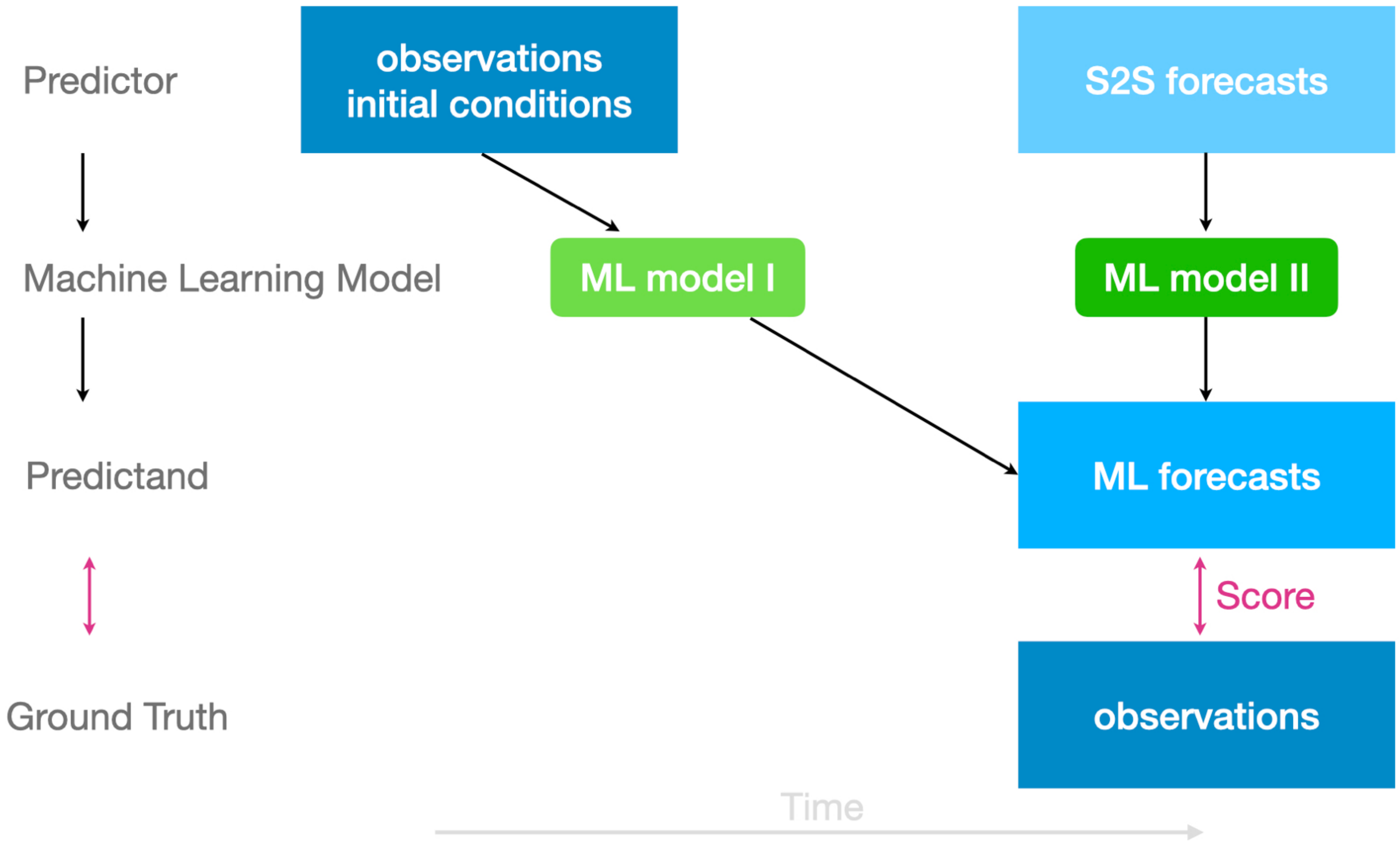
Roberts et al (2020): Impact of Model Resolution on Tropical Cyclone Simulation Using the HighResMIP–PRIMAVERA Multimodel Ensemble

Low-resolution models reproduce only a fraction of the observed cyclones, and are not able to reproduce intense cyclones

AI and climate predictions

AI to improve climate prediction and projections

- In order to correctly reproduce extreme events climate models need to have a high spatial resolution
- This is not always computationally feasible because of the long integrations required (climate projections) and/or the large number of ensemble members needed to represent the uncertainty range (climate predictions)
- AI algorithms can improve the predictions e.g. exploiting the relationship between the extreme events and its large-scale drivers (which are better predicted by the model).



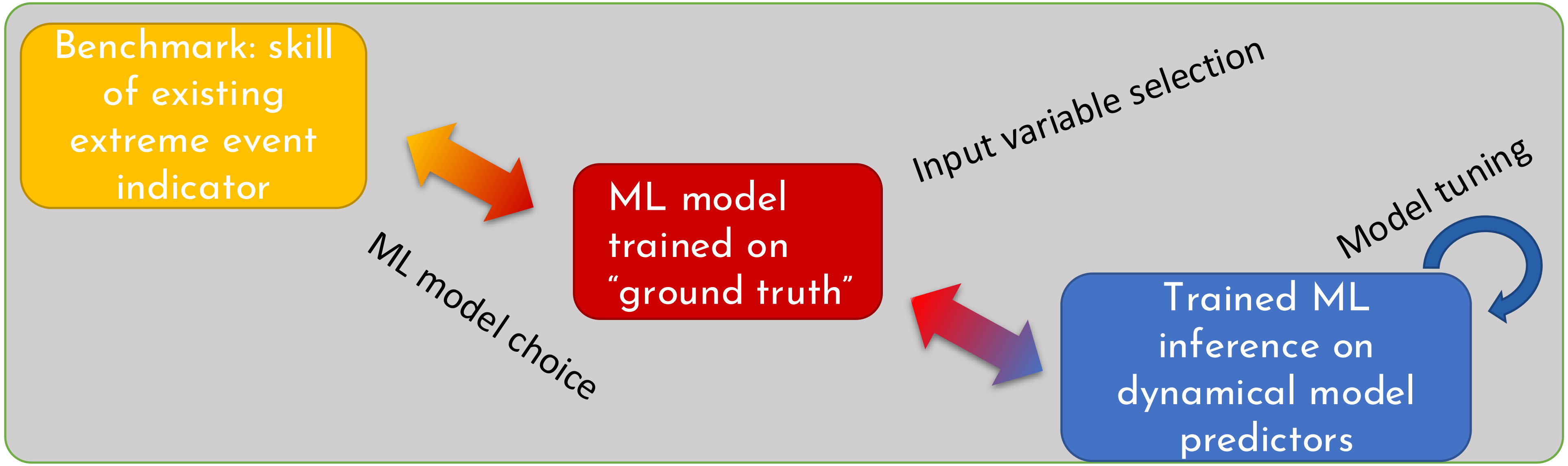
ML model trained on reanalysis to detect the extreme events large-scale drivers, then applied to SPS output

https://s2s-ai-challenge.github.io/

Hybrid AI approach

We choose **a hybrid AI approach** exploiting both existing dynamical forecasts and ML tools:

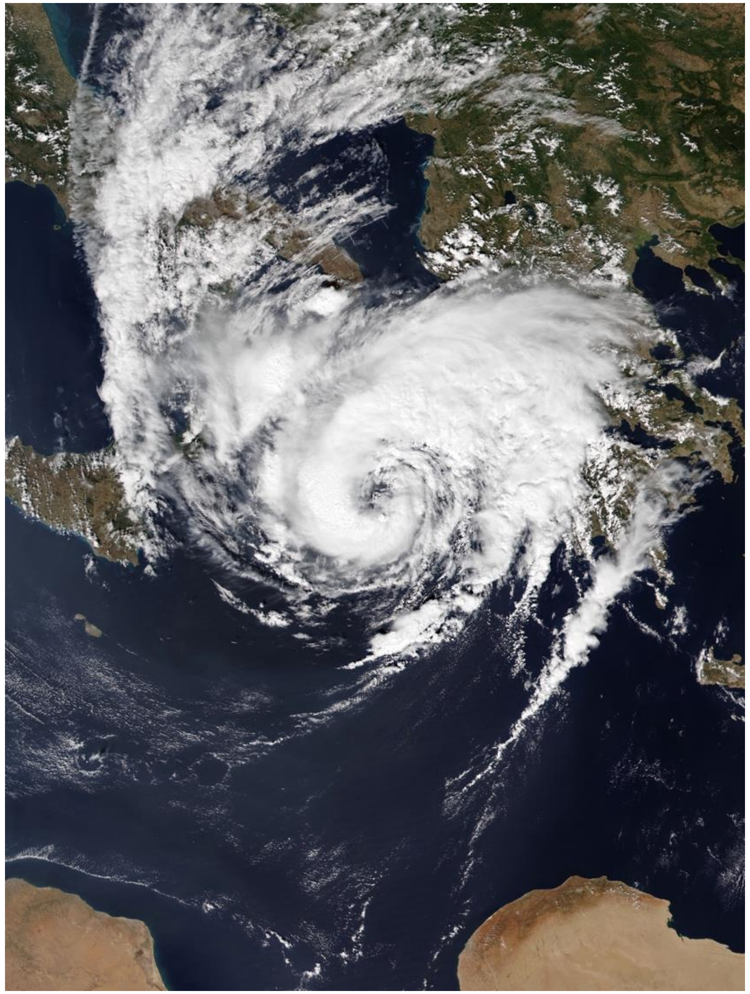
- First the connection between the large-scale variables (predictors) and the extreme of interest (predictand) is established in the "ground truth" by training ML models on observational/reanalysis dataset
- The trained ML model is then applied in inference mode on the large-scale predictors from the dynamical seasonal forecast model hindcasts, and the prediction compared with observations.
- The ML model is tuned to compensate the effect of the dynamical model bias on the predictive skill.



Mediterranean cyclones and climate prediction

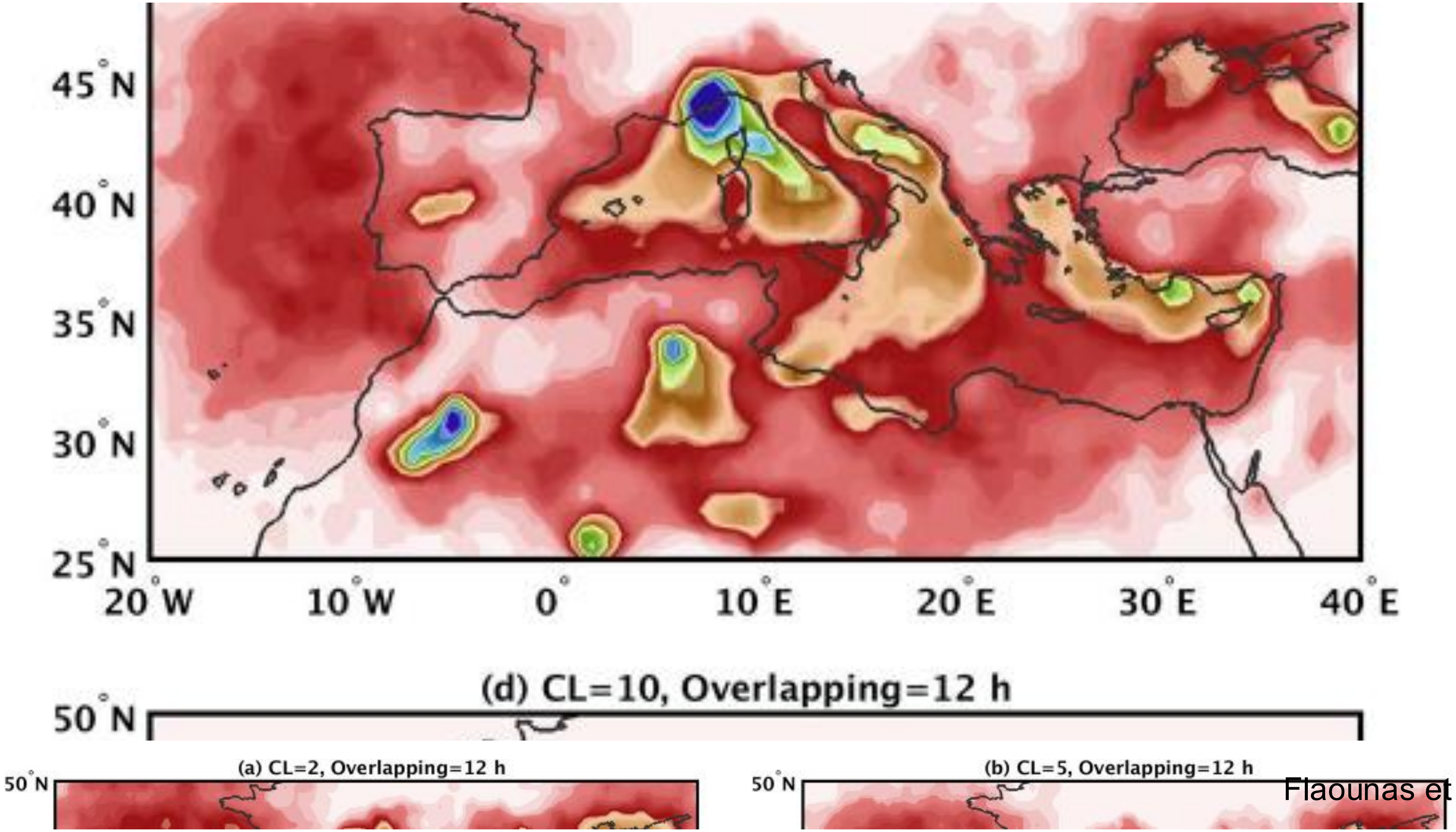
Background

- Cyclones form frequently in the Mediterranean basin due to region location and the complex topography
- Even though smaller and shorter lived than cyclones forming in other ocean basins Med cyclones cause severe damage in the highly populated coasts of the region
- A number of different dynamical mechanisms for cyclone genesis and intensification play a role, resulting in the occurrence of different types of low-pressure systems, ranging from mid-latitude to tropical-like cyclones



Ianos (September 2020)

Med cyclones climatology

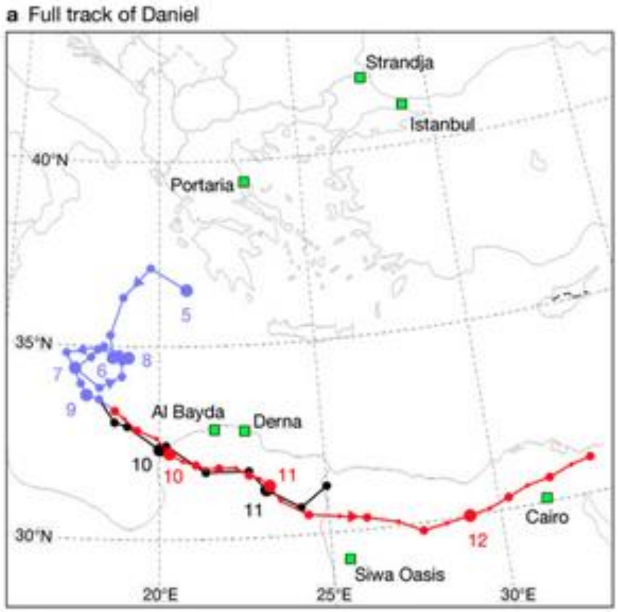


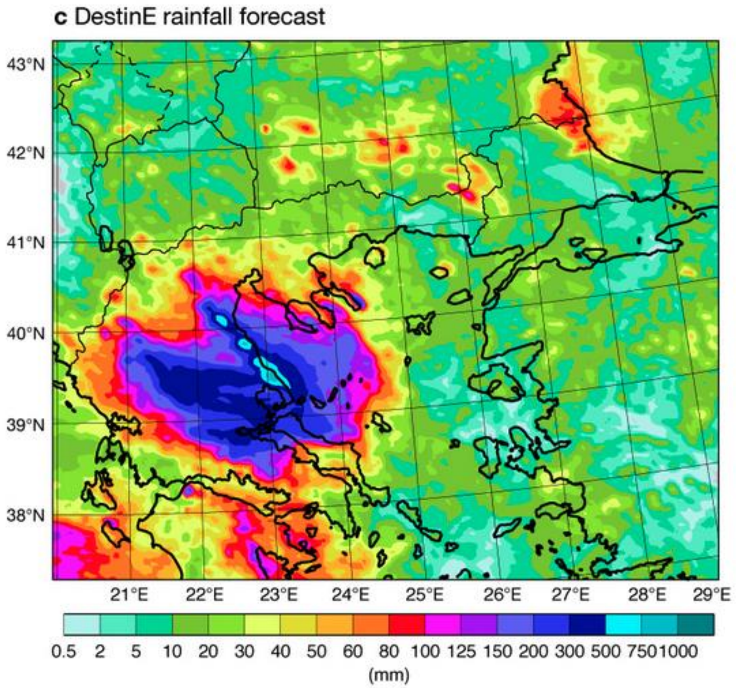
Flaounas et al. 2023

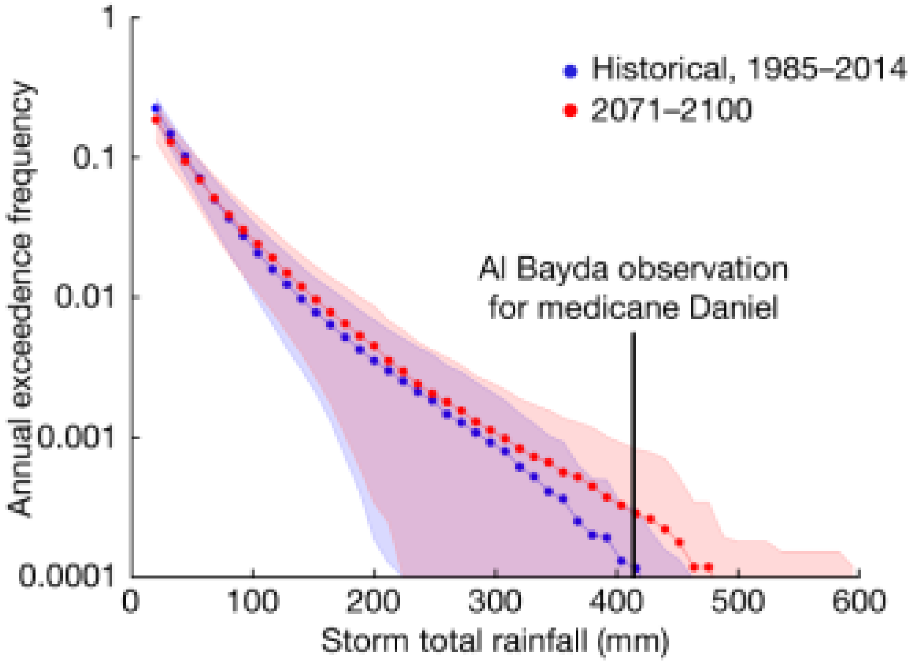
Figure 13. Spatial density of cyclone tracks for composite tracks of four different confidence levels (CLs) and an overlapping criterion of 12 h. Spatial density expresses the average count of track points per year within a circular area with a radius of 0.5°.

Storm Daniel

- *Storm Daniel in September 2023 was the costliest cyclone outside the North Atlantic (> 20 B US$)*
- *The deadliest cyclone globally since 2013 (10.000 fatalities estimated)*








© Halil Fidan/AA/picture alliance

manuscript submitted to *JGR: Machine Learning and Computation*

Understanding and Modeling Mediterranean Cyclone Activity Using Machine Learning

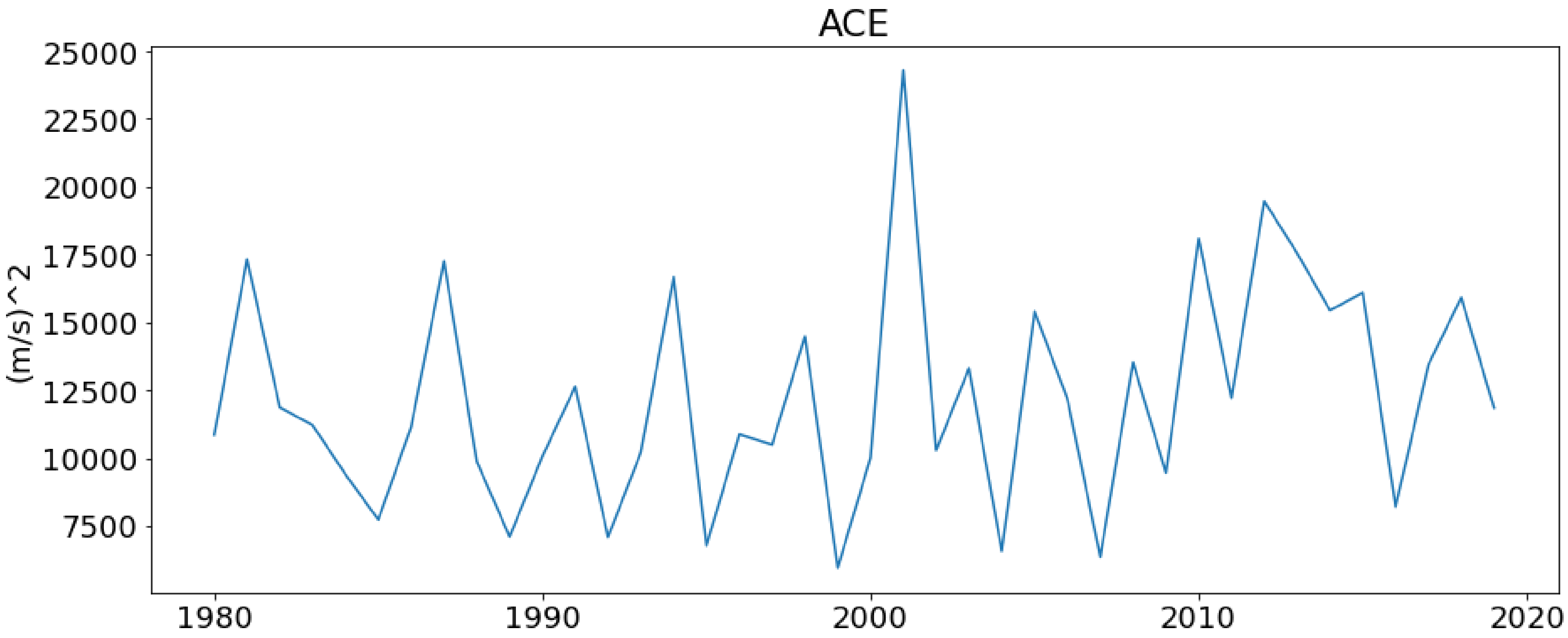
Guido Ascenso[1,2], Luca Proserpio[1], Leone Cavicchia[2], Enrico Scoccimarro[2], Matteo Giuliani[1], and Andrea Castelletti[1,2]

Target

Common metrics used for cyclone activity include cyclone number, cyclone days and ACE (accumulated cyclone activity).

Here we focus on ACE, which has a number of advantages:

- Naturally gives more weight to more intense cyclones, with no need to impose ad hoc filters
- Less sensitivity on the details of the cyclone detection scheme used to produce the ground truth database

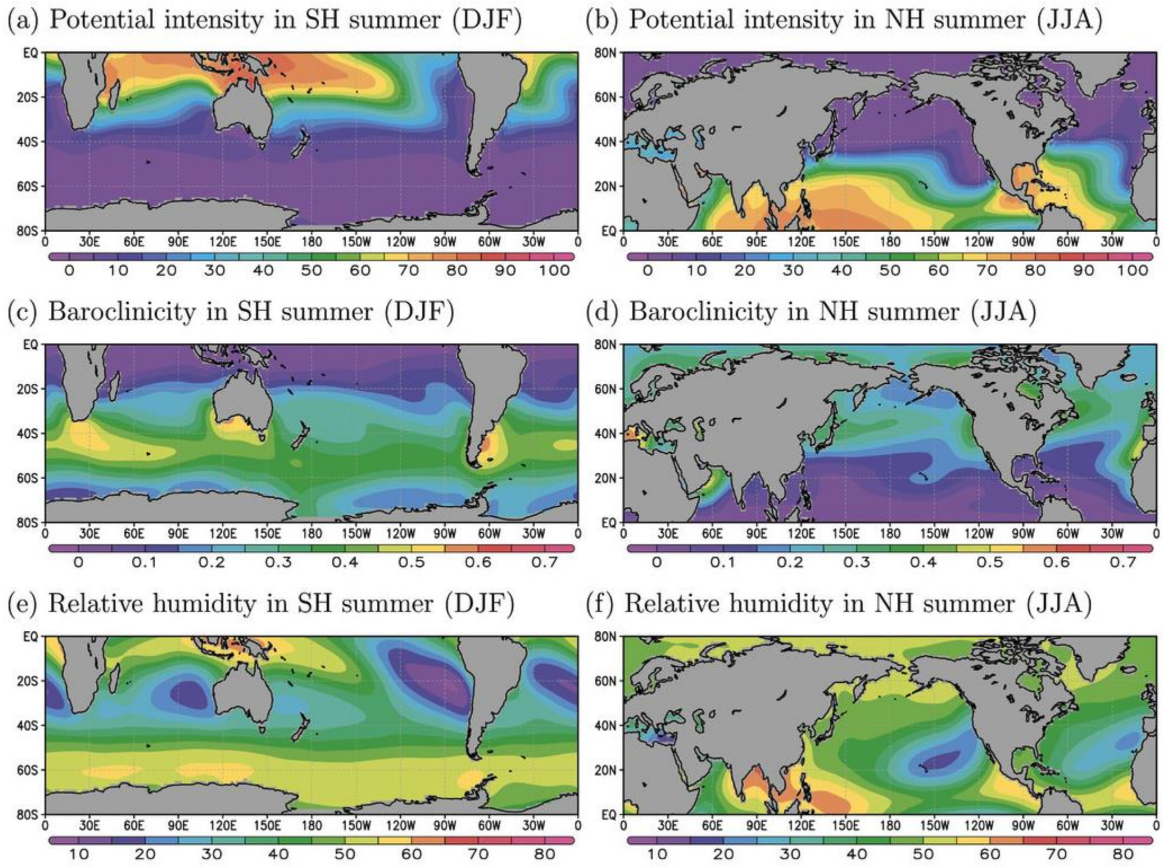


$$\mathrm{ACE} = 10^{-4} \sum v_{\max}^2$$

Drivers

Feature	Acronym	Units	Link to cyclones
Sea surface temperature	SST	K	Warmer waters provide latent heat necessary for cyclone intensification (Pytharoulis et al., 2018).
Absolute vorticity at 850 hPa	AV	s^{-1}	Areas of positive absolute vorticity are associated with the development of low-pressure systems (Miglietta & Rotunno, 2019).
Relative humidity at 600 hPA	RH	%	High-altitude moisture supports deep convection, which is critical for cyclone formation and persistence (Messmer et al., 2017).
Vertical wind shear	V_{shear}	m/s	Weak shear favors tropical cyclogenesis, whereas strong shear disrupts the organization of convection (Miglietta & Rotunno, 2019).
Eady Growth Rate	EGR	s^{-1}	Indicative of baroclinic instability, and therefore particularly useful in the context of mid-latitude cyclones (Besson et al., 2021).
Mean sea-level pressure	MSL	hPa	The gradient between the sea-level pressure at the center of the storm and that of its surroundings is the main driver of the strong horizontal winds towards the center (von Storch et al., 2024).
Air density over the ocean	ADO	km/m^3	For tropical-like storms, lower air density translates into higher efficiency and more fuel to the storm (Abdalla & Cavaleri, 2002).
Convective available potential energy	CAPE	J/kg	The CAPE is a measure of the amount of energy available for convection, and it is correlated with the highest possible vertical speed inside an updraft.
Convective velocity over the ocean	CVO	m/s	An indicator of vertical air motion over water bodies (Dafis et al., 2020).

Table 1. Pool of candidate features for predicting ACE.



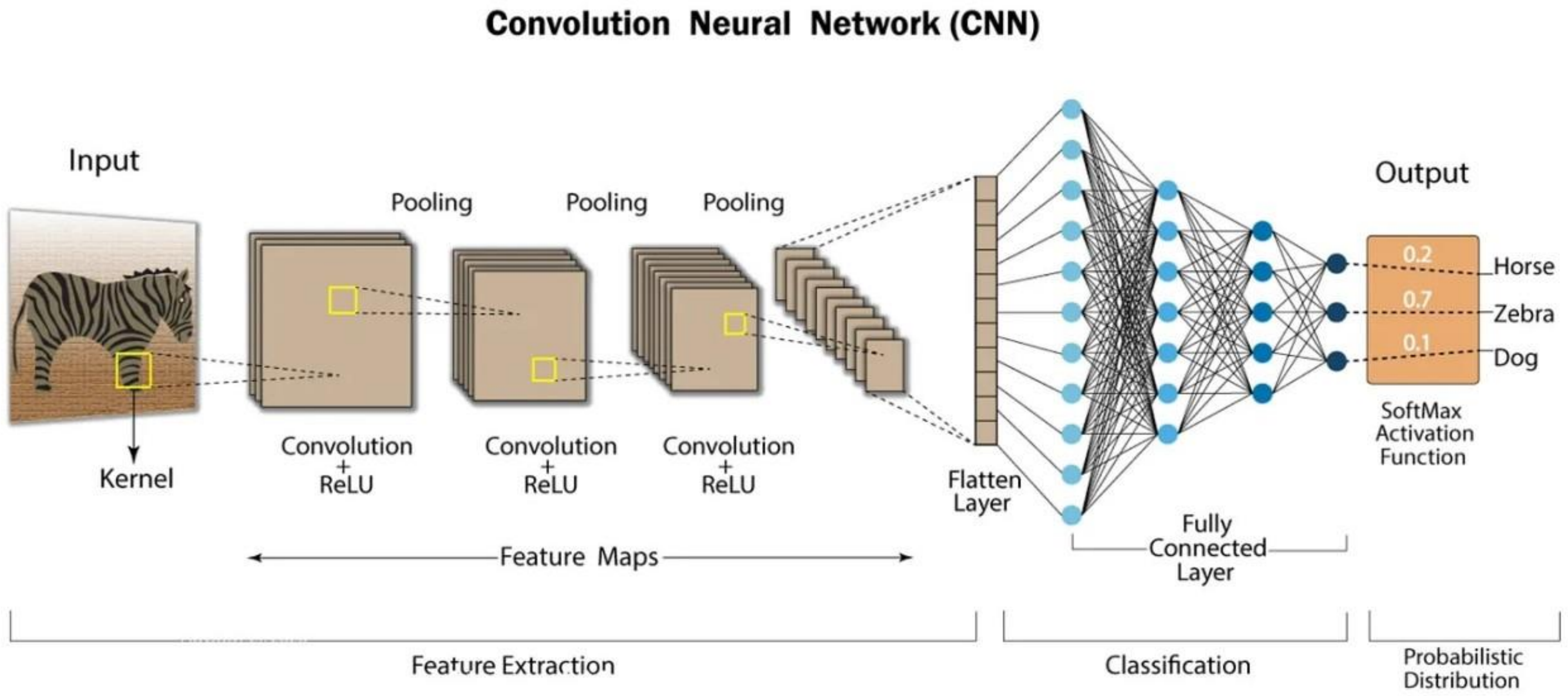
From Yanase et al. 2014

ML models

Different ML models tested to check whether some of them look more fit for purpose.

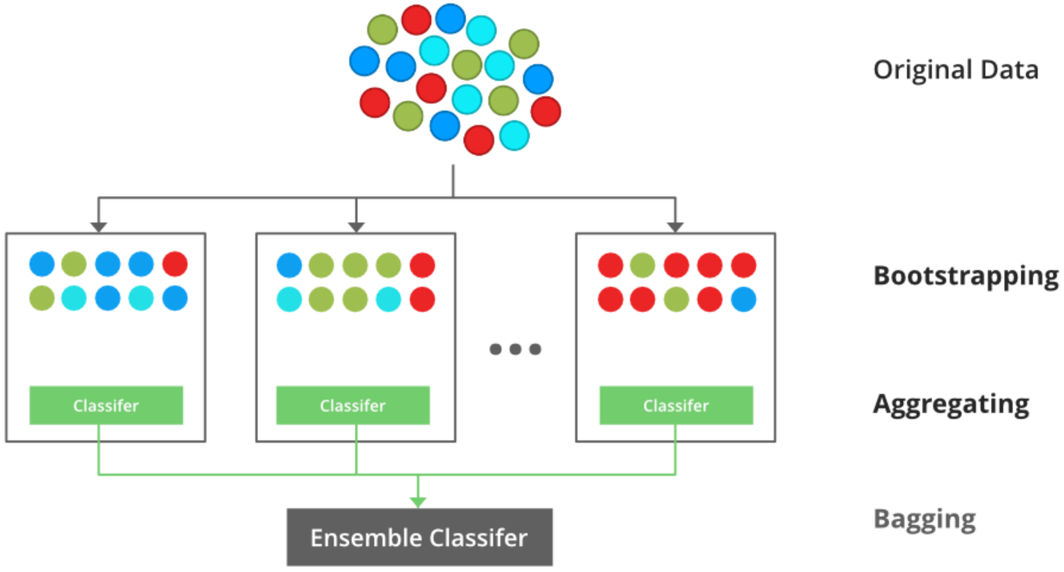
According to the dimensionality of the problem choice of model with intermediate number of parameters.

- First model is a CNN, takes maps of atmospheric fields as input

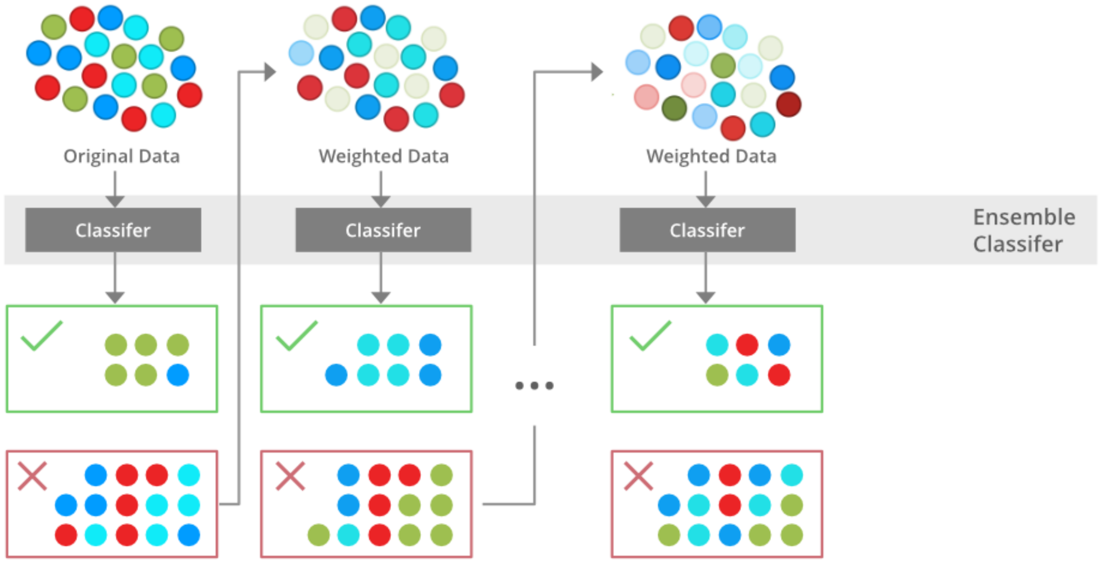


ML models

- Model 2 and 3 are based on tabular data, implementing two different *ensemble learning* approaches: bagging and boosting;
- In the bagging approach, several weak learners are trained in parallel, then aggregated (e.g. majority voting)
- In the boosting approach weak learners are trained sequentially by assigning weights to the sample

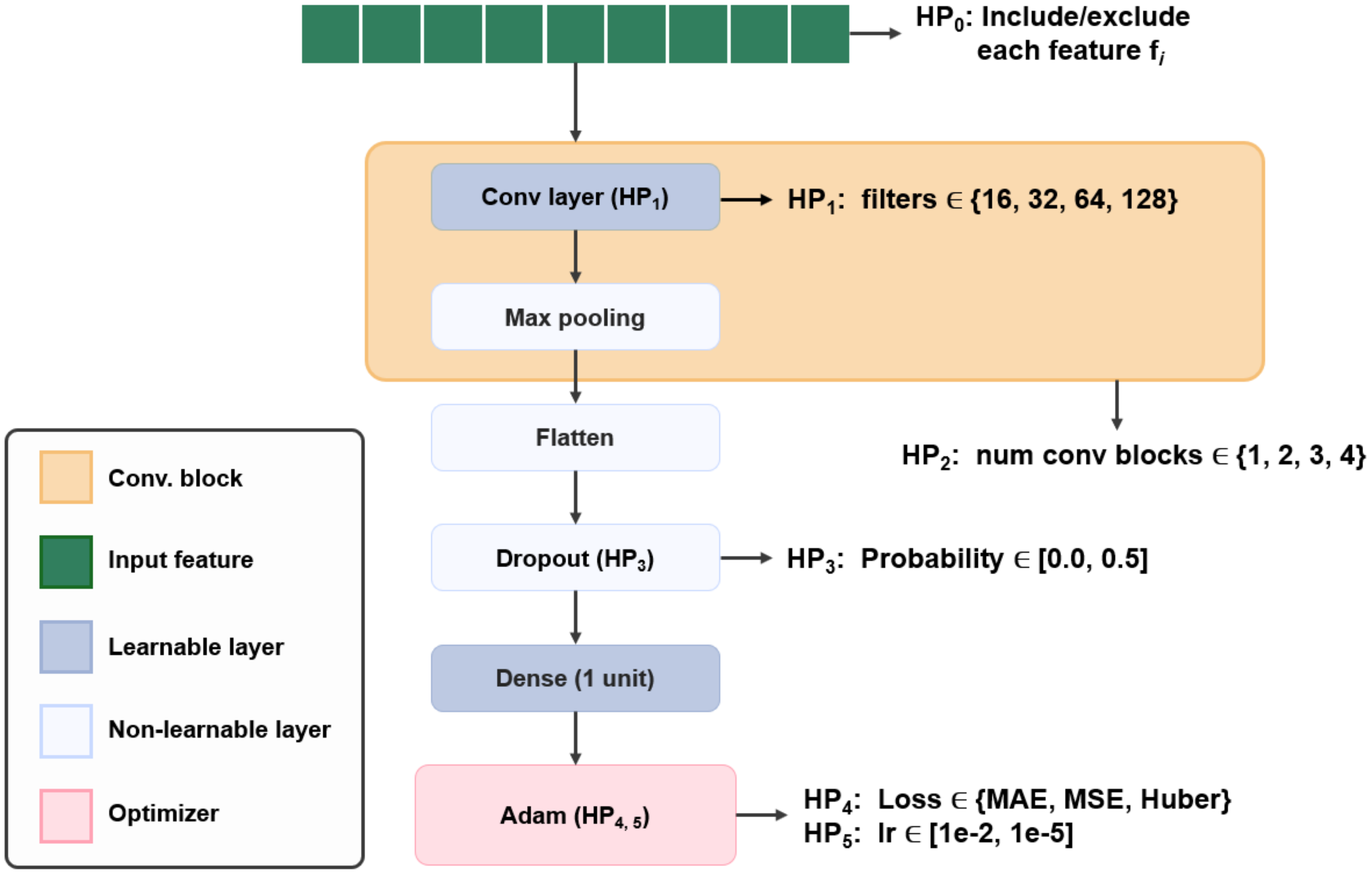


Random forest

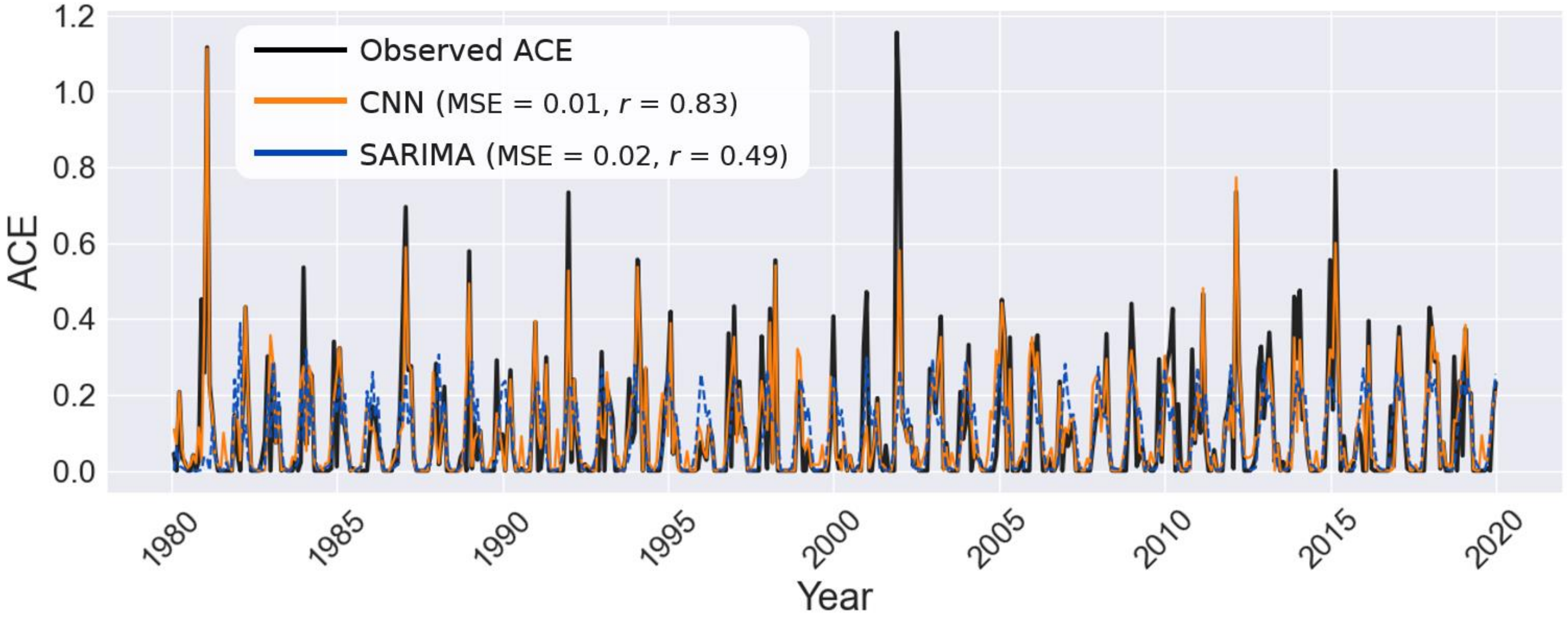


XGBOOST

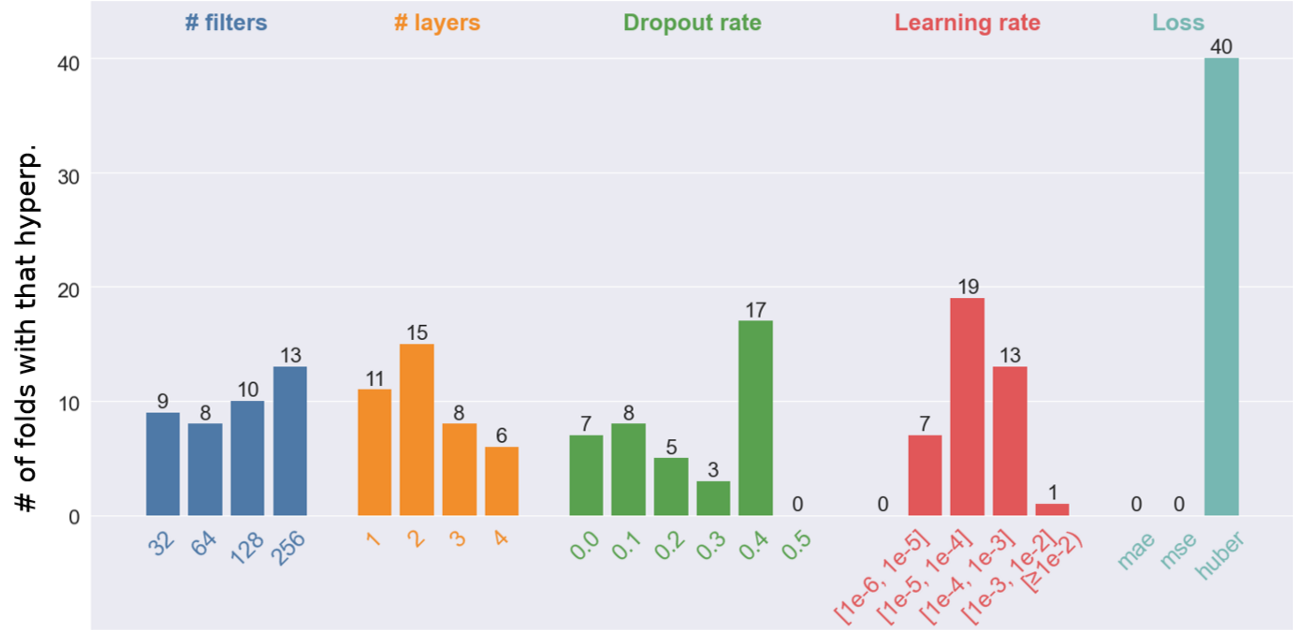
Architecture & hyperparameter tuning



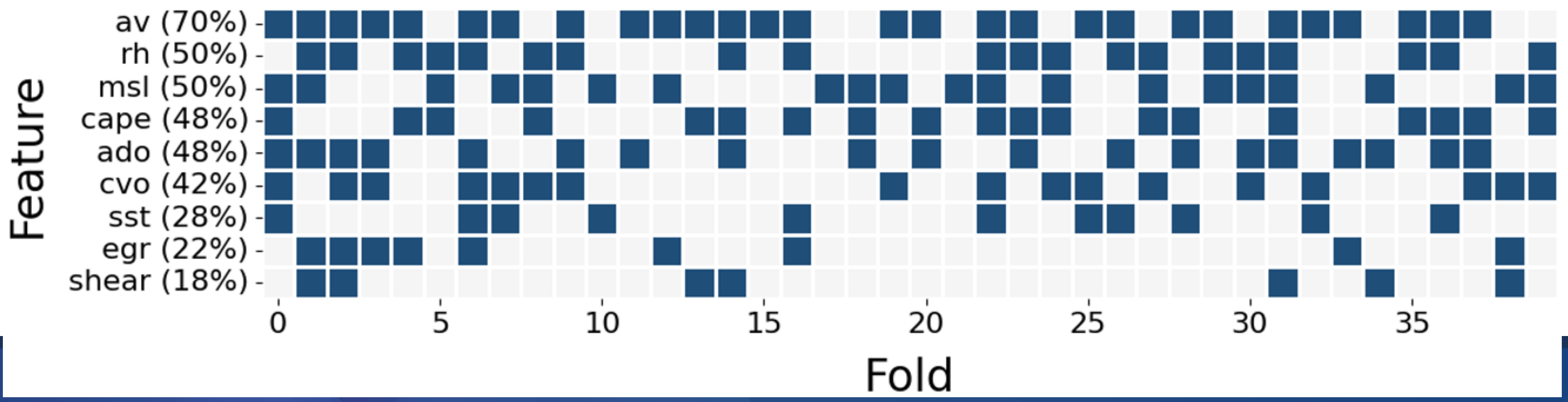
Results: model skill



Results: parameters and feature selection



Feature Selection Across Outer Folds



cmcc
Centro Euro-Mediterraneo
sui Cambiamenti Climatici

www.cmcc.it